PartNet: A Large-Scale Benchmark for Fine-Grained and Hierarchical Part-Level 3D Object Understanding

Notes and Relevance for Game Affordance Mapping Project

Source:
Mo, K., Zhu, S., Chang, A. X., Yi, L., Tripathi, S., Guibas, L. J., & Su, H. (2019). PartNet: A Large-scale Benchmark for Fine-grained and Hierarchical Part-level 3D Object Understanding. CVPR 2019.

---

1. Overview

PartNet is a large-scale dataset of 3D objects annotated with detailed, instance-level, and hierarchical part information.

It enables fine-grained geometric analysis, object understanding, and data-driven reasoning about how objects are composed and function.

Dataset Statistics:

- **Total Models:** 26,671 3D shapes
- **Part Instances:** > 573,000
- **Categories Covered:** 24 (e.g., chair, table, bed, lamp, door, cabinet, bag, etc.)
- **Base Source:** Built on top of ShapeNetCore, extending it with detailed semantic part annotations

---

2. Purpose and Contribution

PartNet was developed to address a gap in 3D vision research: most existing datasets had object-level labels but lacked consistent, fine-grained, and hierarchical part-level annotations.

It supports the study of:

- Fine-grained shape analysis
- Hierarchical semantic reasoning

- Dynamic scene understanding
- **Affordance and functionality learning** (relevant to this project)
- 3D shape generation and simulation

---

3. Dataset Structure

Each object model in PartNet is represented with:

3D Mesh Geometry

- Full-resolution mesh from ShapeNetCore
- Stored in .obj format (triangular faces)

Hierarchical Part Annotations

- Labels structured in a tree hierarchy (e.g., chair → backrest → slat)
- Enables learning at multiple semantic levels

Instance-level Segmentation

- Distinguishes between repeated components (e.g., 4 legs of a chair)
- Facilitates instance-aware learning

Fine-grained Semantic Labels

- Consistent naming across models (e.g., "seat," "handle," "drawer front")
- Supports high-level reasoning about functionality and design

---

4. Benchmarking Tasks

The authors propose three benchmark tasks to evaluate 3D part-level understanding:

Task	Goal	Output Type
Fine-Grained Semantic Segmentation	Assign each point or face a fine-grained part label (e.g., seat, handle, drawer)	Per-point or per-face labels
Hierarchical Semantic Segmentation	Predict multi-level part labels consistent with the annotation hierarchy	Hierarchical label tree

Instance Segmentation	Identify distinct instances of repeated parts (e.g., four chair legs)	Instance IDs per part

---

5. Proposed Baseline Method: PointNet++

The original paper establishes baseline performance using **PointNet++**, a hierarchical neural network that processes point clouds rather than meshes.

Each 3D object mesh is sampled into a point cloud (~10k–50k points), where each point inherits its part label.

Key Features of PointNet++:

- Operates directly on unordered point sets
- Learns local geometric features using hierarchical sampling and grouping
- Robust to input permutation and varying density
- Efficient for large-scale datasets like PartNet

Relevance: This is directly relevant to the Game Affordance Net hybrid pipeline, which uses **DGCNN (Dynamic Graph CNN)**—a successor model that builds on the PointNet++ philosophy but adds dynamic local graph learning for improved shape understanding.

---

6. Dataset Hierarchies and Label Example

Each object category has a predefined semantic hierarchy.

Example: Chair

```
Chair
├── Seat
├── Back
│   ├── Backrest
│   └── Armrest
├── Legs
│   ├── Leg Front Left
│   ├── Leg Front Right
│   ├── Leg Back Left
│   └── Leg Back Right
└── Support Bars
```

This structure allows models to reason about:

- **Object composition** (how parts fit together)
- **Functional affordances** (how those parts relate to human actions)

---

7. Technical Details

Data Format:

Each object includes:

- `.obj` mesh file
- `.json` annotation tree with part hierarchy
- `.pts` or `.ply` sampled point cloud with per-point labels

Label Schema:

- 337 fine-grained part categories organized across 24 parent object types

Sampling Strategy:

- Uniformly sampled 3D points from mesh surfaces, maintaining consistent coverage for dense labeling

Annotation Process:

- Human annotators manually labeled parts through a web-based 3D annotation interface, ensuring high accuracy and consistency

---

8. Dataset Usage in This Project

Integration with Hybrid Point–Mesh Pipeline

For the affordance learning system, PartNet provides both the geometry (mesh) and semantic part labels needed to train supervised models.

Training Data Preparation Workflow:

1. Load the PartNet mesh for each object

2. Sample 20–50k surface points
3. Inherit the part label from the original annotations
4. Map part → affordance (e.g., seat → sit, handle → grasp, shelf → support)
5. Train DGCNN for per-point affordance segmentation
6. Project predicted affordance labels back onto the mesh for Unreal integration

This enables the model to learn **shape-to-function mapping** — understanding how geometry implies possible actions.

---

9. Why PartNet Is Suitable

Criterion	Relevance
Large Scale	Over 500k annotated parts across 20k+ shapes ensures generalization
Fine-Grained Detail	Essential for detecting affordances that depend on small geometric variations (e.g., handle vs. rim)
Hierarchical Labels	Supports multi-level affordance reasoning — local vs. global object function
Consistency	Standardized annotation schema across categories supports cross-object learning
Public Benchmark	Widely used in recent research (PartAfford, SegAffordSplat, MeshPointNet), ensuring compatibility with state-of-the-art pipelines

---

10. Relevance to Affordance Research

PartNet is not an affordance dataset by itself—it provides **structural and semantic part annotations** that can be mapped into affordance categories.

Example Mapping:

Part Label	Mapped Affordance
Seat	Sit
Handle	Grasp
Drawer	Pull

Basin	Contain
Shelf	Support
Button	Press

This mapping acts as the **supervised training signal** for affordance detection models such as DGCNN.

---

11. Related Datasets and Extensions

Dataset	Focus	Relation to PartNet
PartNet-Mobility	Adds motion and articulation (e.g., open/close)	Useful for dynamic affordances
ShapeNetCore	Object-level 3D models	Parent dataset
PartAfford	Maps PartNet labels to affordance labels	Directly relevant
Shape2Motion	Adds kinematic information	Enables learning of movable parts

---

12. Summary

- PartNet provides the **semantic and geometric foundation** for affordance learning
- It enables training data creation through **part-to-affordance mapping**
- Supports **dense point cloud segmentation**
- Serves as the benchmark for evaluating 3D understanding
- Its integration into the **hybrid pipeline** (point-based DGCNN → mesh projection → Unreal tag import) directly supports the automation of geometry-aware affordance detection in 3D game environments